

A Study The Review Of Video Encoding And Video Compression

Maya Chowksey^{1*}, Dr. Ravindra Tiwari²

¹ Research Scholar, LNCT University, Bhopal

² Associate Prof, LNCT University, Bhopal

Abstract- The term "video compression" refers to the process of encoding a video file in a way that results in a smaller file size. It can be difficult to record a lengthy video sequence due to the size of the resulting video files. When a video is compressed, its file size is decreased because unnecessary information is removed. Compressing a video involves a lengthy procedure in which the original files are reduced in size significantly. Video coding is what establishes the efficiency of compression on live-action entertainment videos & collaborative mobile communication situation needed to access the little video frames of predetermined dimensions. The quantity of data needed to depict a video frame is drastically decreased without a noticeable drop in image quality. Most commonly used video encoding protocols, including H.264 and H.263, compress video frames before sending them over the network. This study explicitly gives an insight into the theoretical aspects and background of video encoding and video compression.

Keywords- video compression, video encoding, encryption,

-----X-----

INTRODUCTION

Recent developments in Multimedia Networking & Coding have broadened the field's potential applications to include, for example, the reliable transmission of real-time audio, video, & image streams. The increasing popularity of online video sharing platforms like YouTube and Vimeo is fueling the need to send encoded media files across wireless networks. Online video hosting sites such as You Tube & Vimeo. While doing so, it has also gone out of its way to help users meet the growing need for high-quality online video streaming with little lag time between sender & receiver. People need a new form of communication that combines data, images, audio, and video because as communication networks build, the various departments of an enterprise exchange information more closely through networks, leaving people dissatisfied with telephones, fax machines, e-mail, and other traditional voice & text communications. To increase network security, video must be encrypted before transmission (Hua-Zhen Yao et al. 2010). For video encryption to work, a combination of cryptography & multimedia technologies is required. Encryption is a method for preventing unauthorized individuals from viewing or altering sensitive data, such as files, photos, or financial transactions conducted online. The plaintext data is transformed into an unintelligible form utilizing a mathematical function called a cipher & key used only by the intended recipient (cipher text). The key makes the encrypted information really one-of-a-kind,

whereas the cipher is the generic recipe for encryption. It can be deciphered by authorized users only if they have both the correct key & same cipher. Common methods of authentication include passwords, tokens, & biometrics; keys are typically long sequences of integers (like the fingerprint).

FUNDAMENTALS OF VIDEO CODING

The purpose of this article is to present a high-level conceptual overview of numerous components and their interaction scenario by briefly discussing the core notion of Video Coding systems. In addition, the procedure (Tool Chain) as a whole comprises a collection of auxiliary systems, which stretch from the first Acquiring of the video frames & subsequent Transmission to the final Display. In order to evaluate video compression, the study primarily focuses on standard encoding & decoding procedures and their respective analyses. The video coding system block diagram is shown in Figure 1.

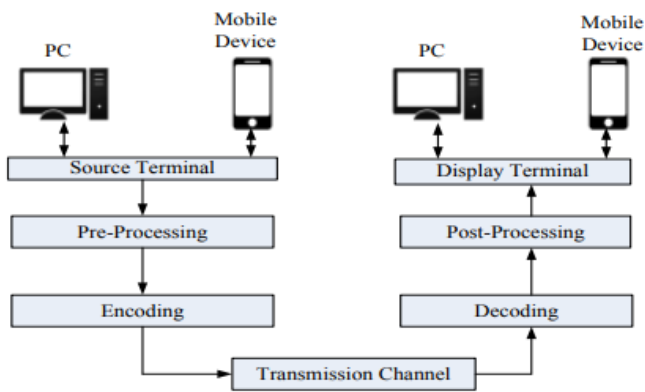


Figure 1 Block diagram of Video Coding Systems

An overview of the entire video coding system is shown in Figure 1; below, we outline each block and how it depends on the others to process a video frame sequence.

- **Video Acquisition**– The process of acquiring & processing video is crucial to digitizing the original video sequence in a way that is independent of time. Capturing natural landscapes, scanning photographic film materials, and evaluating the signal components to disclose their unique properties are all possible thanks to this method [Iain E 2004].
- **Pre-Processing**– Raw uncompressed video data sequences undergo this procedure, which describes in detail the trimming, color format conversion, & correlation steps taken in their execution [David R 2014].
- **Encoding Process**– Video frames are converted into a coded bit stream sequence as part of the procedure. Regardless of the presence or absence of noisy data bits, it also provides a concise representation of the video stream. However, the compression ratio is improved by the encoding, which makes the encoded video stream better suited for a particular transmission method [Y-H Chen 2015].
- **Transmission**– It makes it possible to encapsulate a digital bit stream by taking into account the proper packet format. To maximize throughput and packet delivery ratio, the data packets are then sent through a transmission channel. This includes not just the aforementioned case but also loss protection & loss recovery during the transmission of bits from an encoded video stream [T. Nguyen 2015].
- **Decoding Process**–The decoder takes in the full bit stream and puts it into a buffer for later use. The normative techniques, provided under standard coding specification, for decoding require further conversion of the incoming encoded bit stream data into a video sequence. To accommodate the needs of the transmission, the compression or video encoding is lossy (Data rate & Bandwidth

constraints). Whereas Decoding procedure reveals substantial data loss from the video stream & results in a close approximation of the actual source video stream, the latter is not always accurate. Concealment tactics are used by the decoder to restore the damaged video sequence & make it look as near as possible to the original [S. Panayides 2015].

- **Post Processing**– It is the responsibility of the post-processing activities to make use of the video frame sequence that was produced as an output during the decoding phase. The focus is on improving the image quality by rebuilding the movie frame by frame.
- **Display**– To transform the reconstructed video frames into a format suitable for color presentation in a display device. The screen serves as a conduit between the computer and the human eye. The output timing of a video sequence's still images can be manipulated to create the desired effect [L. Trzcianowski 2015].

FUNDAMENTALS OF VIDEO COMPRESSION

The history of video compression and its potential uses in new technologies are discussed in detail here. To create a coded video stream, video compression removes unnecessary data from an original video sequence. The technique also includes maximizing the efficiency with which a coded video stream is stored and sent, i.e. the transmission of a video sequence without regard to storage or bandwidth constraints. Video encoding utilizes compression algorithms on the bit stream sequence before transmission & storage. Encoding delay refers to the amount of time it takes for the compression method to remove the superfluous data. Once the video has been encoded, it must be sent through a communication channel and run through a decoder, de-compressor, or other ancillary components before it can be reproduced as faithfully as the original raw video. To be more specific, compression and decompression are two of the processes that make up a codec's core functionality. It is more probable to happen, that different compression mechanisms provide varied outcomes in terms of bit rates and quality levels.

In order to lower the bit rates of streaming movies, various compression standards have been developed, some of which include ISO/IEC MPEG, MPEG-2, AVC, HEVC & ITU-T VCEG. Even though AVC standards are more cost-effective than traditional MPEG at keeping video quality stable, they are nonetheless extremely pricy (between storage and bandwidth constraints). However, compared to AVC and MPEG-2, HEVC's efficient video compression reduces file sizes by up to 50% and 75%, respectively. ISO/IEC MPEG and the Video Coding Experts Group of the International Telecommunications Union (ITU-T) collaborated to

create the HEVC (H.265) standard [G. He 2015; Y-J Ahn 2014].

When combined with the cutting-edge software & hardware frameworks currently available on the market, the HEVC offers improved usability and performance. As a result, HEVC has a greater chance of providing significant and improved compression of high-end definition videos, such as High Definition (HD), Full High Definition (FHD), & Ultra High Definition (UHD), while yet maintaining a low bit rate.

Table 1 History of International Video Compression Standards

Year	Standard	Publisher	Popular Implementations
1988	H.261	ITU-T	Videoconferencing, Video Telephony
1993	MPEG-1 Part 2	ISO, IEC	Video-CD
1995	H.262/MPEG-2 Part 2	ISO, IEC, ITU-T	DVD, Digital Video Broadcasting, SVCD
1996	H.263	ITU-T	Videoconferencing, Video Telephony, Video on Mobile Phones (3G Phone)
1999	MPEG-4 Part 2	ISO, IEC	Video on Demand
2002	H.264/MPEG-4 AVC	ISO, IEC, ITU-T	Digital Video Broadcasting, iPod Video, Apple TV, HD DVD
2008	VC-2 (Dirac)	ISO, BBC	Video on Demand, HDTV broadcast, UHDTV

Data compression methods fall into two categories: (i) predictive coding and (ii) transform coding. Predictive coding exploits the redundancy in the data. As the name indicates, transform coding, transforms the input video frame into different arrays and packs them into a number of small samples. Other compression algorithms are either generalizations or combinations of predictive coding and transform coding. It is likely that video compression results in some distortion of data because of analog to digital conversion and omission of some insignificant information. However, efficient video compression techniques try to reduce distortion. The timeline of video compression techniques and standards are tabulated in Table 1.

Video Encryption

Video encryption is the process of converting a plain video into a cipher video by using a key or hash function which makes it unreadable. The reverse process of video encryption is called video decryption which converts the unreadable cipher video into readable plain video using the key or hash function which has been authenticated during encryption process.

Basics of Video Data

Video is a live stream which can be recorded, played or displayed by information processing devices and

real time data processing units, such as computerized electronic devices (Gordon 2003). The video has been classified into two categories: (i) linear and (ii) non-linear. The linear video is an active content stream which has no navigation control. Cinema presentation is one of the best examples for linear video. The non-linear video stream has use interactive feature where it supports all controls during navigation. Computer games, computer based learning and hypermedia are examples of non-linear video. Video presentations are live streams or recorded file. A recorded video application can interact with the user through its navigation system, whereas, a live video stream can interact with the user directly without any interface. Video is a sequence of frames and its speed can be measured by using *frame rate* (Matthias 2008). Number of frame images per unit time is called frame rate of the video. The frame rate of old mechanical camera ranges from 6 or 8 frames per second and the frame rate of modern digital cameras ranges from 120 to more frames per second. The Phase Alternating Line (PAL) and Sequential Color with Memory (SECAM) standards specify that frame rate of video camera should be 25 frames/sec and National Television System Committee (NTSC) specifies 29.97 frame/sec. Film presentation is one of the interactive video applications which has lower frame rate of 24 frames/sec. To get the illusion from video frame, the application should be projected with a minimum frame rate of 15 frames/sec.

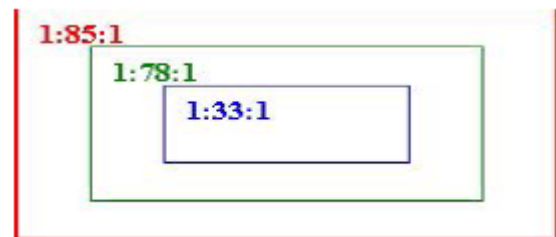


Figure 2 Aspect Ratio of Traditional Television in RGB

Depending on the device, different dimensions are used for color and gray scale video signals. Aspect Ratio (AR) is used to measure dimensions of the video screen. It is the ratio of screen width and height. Figure 2 shows the Aspect Ratio (AR) of traditional television. The aspect ratio of a traditional television screen is 1:33:1 and High Definition (HD) television is 1:78:1.

The color representation of a video is called its color model. There are various types of color models used in video display system. In NTSC and PAL standards YIQ color model which is similar to YUV color model is used whereas in SECAM standard, YDbDr color model is used. Color representation of UV color model is shown in Figure 3.

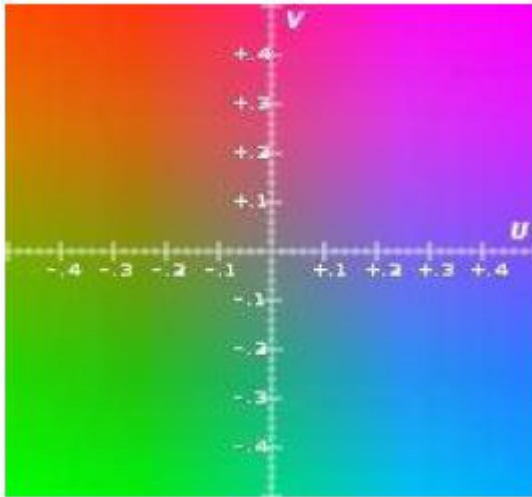


Figure 3 U-V Color Plane

RGB and YUV Representation of Video Signals

In general, the Red, Green, Blue (RGB) color system is used in representation of color images and video (Iain 2003). Appropriate contribution of RGB can generate any desired color. In general, a color is represented using luminance and chrominance. Let Y be the luminance which represents the brightness of the color. Luminance can be calculated by using the weighted sum of three colors R , G and B . Color difference of these color models Cr , Cg and Cb can be calculated by subtracting the luminance from each primary component using the following expressions:

$$Cr = Wr * (R - Y) \tag{1}$$

$$Cg = Wg * (G - Y) \tag{2}$$

$$Cb = Wb * (B - Y) \tag{3}$$

Where, Wr , Wg and Wb are weights of R , G and B . Among these three color differences Cr , and Cg are linearly independent. The color difference Cb can be expressed as a linear combination of the two linearly independent signals R and G . Luminance Y and any of the two color differences are used to represent a color. Some of the popular standards such as NTSC, PAL and SECAM use three components, luminance Y , blue color difference U and red color difference V to represent a color (Iain 2003). This is called YUV color system. The RGB color signal can be converted into YUV model by using the following expressions:

$$Y = (0.257 * R) + (0.504 * G) + (0.098 * B) + 16 \tag{4}$$

$$U = (0.439 * R) + (0.368 * G) - (0.071 * B) + 128 \tag{5}$$

$$V = -(0.148 * R) - (0.291 * G) + (0.439 * B) + 128 \tag{6}$$

Human Visual System (HVS) is less sensitive to chrominance than brightness. This is because chrominance signals are represented by lower

resolution than luminance. Therefore, YUV is preferred over RGB in some standards. Video data processing quality is very important parameter and is used to measure signal strength of video signal. Subjective video quality (Ke 1997), (Kotevski 2010), (Maria 2012) (Andrew 1999) for video processing system is evaluated as follows:

- a) Select all video sequences for testing
- b) Define a system for evaluation
- c) Define a method for presenting video sequences to experts
- d) Invite enough number of experts
- e) Do testing on video
- f) Collect expert's ratings
- g) Calculate average points for each testing based on experts rating

CHALLENGES IN VIDEO ENCRYPTION

Internet users have been experiencing viruses, trojans, hacking and data espionage. The greatest risk for internet users lies in infection of computers with malware and spyware. The economic losses incurred due to officially recorded Internet crimes in Germany amounted to approximately 61.5 million euro dollars in 2010. With the increasing trend in digital communication and multimedia commerce applications, it is important to provide security to audio, video and other information involved in transmission. In addition, the need for video has been growing rapidly in areas such as education, communication, publishing and entertainment in recent years. Combining digital video, database and communication network technologies enhance the ability of delivering vide along with text, images and audio through networks. Applications such as digital libraries, video databases, video-telephony/conferencing, and Videoon-Demand (VoD) are already in use. However, improvement is required in terms of storing huge amount of video data at relatively low cost, efficient organization of video data, retrieving, delivering, and presenting the requested data for easy access for the above mentioned services. One of the major challenges in deploying digital video is the huge volume of uncompressed video which may overwhelm the available communication channels and storage systems.

Two important factors for video data encryption are file size and speed of execution. A video of three hours movie takes nearly 1 GB size. Real time applications such as video conferencing and video telephony require more memory while in operation. It consumes more time to process such large amount of data for storing or transferring through a network. This is due to the large size of video files when compared to text files. Therefore, video data need to be compressed before transmitting or sharing through the network.

Video is the primary commodity in the world of E-Commerce. As technology advances and access to

markets expand, the need to protect video data, to ensure confidentiality, integrity, and availability for making critical personal, business, or government decisions became more important. Without protection of video data, success of an E-Commerce marketplace may be short-lived. Video security plays a major role in providing threat free communication. Implementing weak video security mechanisms can result in the loss of trust, reputation, and money for consumers, businesses, and governments. The sharp increase in the number of fraud, extortion, and identity theft crimes is the primary result of weak video security mechanisms. The cost of implementing basic mechanisms to protect video data is generally much less expensive than a security breach. Weak video security mechanisms can generally be attributed to the lack of understanding and acknowledgment of potential threats to information.

VIDEO COMPRESSION STANDARDS

Several techniques are proposed for video compression. Some of the video compression techniques used for text and image compression show better performance when compared to compressed video data. This is due to large sizes of video files. After compression, the compressed data should maintain its resolution and quality (Andrew 2003). Well known compression techniques are accepted by International Telecommunication Union (ITU) and International Standard Organization (ISO). Some of the popularly used video compression techniques are discussed in subsequent sections.

H.120

The first digital video encoding technique H.120 was introduced in 1984 which was published by International Telegraph and Telephone Consultative Committee (CCITT). In 1988, H.120 was modified with contributions from organizations such as ITU Telecommunication Standardization Sector (ITU-T) and Telecommunication Standardization Bureau (TSB) (Kliarativili 1994).

H.261

ITU Telecommunication Standardization Sector (ITU-T) introduced the first H.26x family video coding standard, H.261 in 1988. H.261 is designed for transmitting video data with data rate in multiples of 64-Kbit/sec over Integrated Services Digital Network (ISDN) lines. But later, this operates at video bit rates between 40-Kbit/sec and 2-Mbit/sec. This standard supports only Common Intermediate Format (CIF) (352x288 Luma with 176x144 Chroma) and Quarter Common Intermediate Format (QCIF) (176x144 Luma with 88x72 Chroma) video frame sizes (Gene 2003). In 1993, new version of H.261 is introduced for processing still picture graphics with 704x576 Luma and 352x288 Chroma resolutions which has backward compatibility. This was the first standard in which the macroblock concept is introduced (Alessandro 1999).

The basic processing unit of H.261 standard is called macroblock. Each macroblock consists of a 16x16 array of luma samples and two corresponding 8x8 arrays of chroma samples, using 4:2:0 sampling and an YCbCr color space. The H.261 standard is helpful in decoding the stored video data (Mee 1998). However, H.261 cannot be used for real time video encoding.

H.263

ITU-T Video Coding Experts Group (VCEG) has developed a low bit rate compression format, H.263 in 1996 for videoconferencing application which is in the H.26x family of ITU-T domain. This standard is used in many real-time applications including sites such as YouTube, Google Video and MySpace (Ahmad 2010). Many websites use this standard for their video encoding (Ujwala 2010). Real Video Codec (RVC) was based on H.263 till Real Video 8 was launched. H.263 is also used in Packet-switched Streaming Services (PSS), Multimedia Messaging Service (MMS) and Internet Multimedia Subsystem (IMS) (Kodituwakku 2010).

H.264

H.264 was introduced in 2002 to mitigate disadvantages of MPEG and to improve the compression performance in broadcasting. The H.264 standard is widely used in satellite and cable TV since it is more convenient than other compression standards (Thomas 2010). It is currently used for video recording, video telephony, video streaming and HDTV streaming over the Internet. When compared to MPEG-2, it provides better video quality with half data rate which is more useful for real-time video data transfer with high speed during video conferencing (Enrico 2006), (Suman 2011). References such as (Iain 2003) and (Thomas 2003) may be referred for additional details on H.264.

MPEG-1

The Motion Pictures Expert Group-1 (MPEG-1) compression standard is designed for compressing low quality video, such as Video Home System (VHS). This can be used for making CDs, in cable TV and digital audio broadcasting (Jurgen 1995). MPEG-1 is a lossy compression standard. Many products and technologies are introduced based on MPEG-1 (Tino 1996), (Gene 1996).

MPEG-2

The MPEG-2 standard comes in three main parts, such as systems, video, and audio. MPEG-2 extends functions provided by MPEG-1 to enable efficient encoding of video and associated audio at a wide range of resolutions and bit rates. Part-1 of MPEG-2 standard specifies two types of multiplexed bitstreams (Kwok 2008). They are program stream

and transport stream. The program stream is analogous to systems part in MPEG-1. It is designed for flexible processing of multiplexed stream and for environments with low error probabilities. The transport stream is constructed in a different way and includes a number of features that are designed to support video communications or storage in environments with significantly higher error probabilities (Chi 2011).

Audio Video Interleave

Audio Video Interleave (AVI) format is introduced as a built-in feature of the Windows Operating System in 1992. Digital file format is used to store the audio and video data. AVI is derived from Resource Interchange File Format (RIFF), which divides a file's data into several blocks (Navas2011).

Dirac

The open source video compression format known as Dirac was introduced by British Broadcasting Corporation Research (BBCR) in 2008, with the help of Schrodinger and Dirac research foundation (Sinzobakwira 2010). This is a high quality video compression standard used in High Definition Television (HDTV) and compete with existing technologies such as H.264 and Video Coding-1 (VC-1). It supports 1920x1080 resolution (Chi 2011), (Patrick 2009), (Vishesh 2011) and more data rate like MPEG-4 Part 2 and MPEG-2 Part 2. The motion compensation and interframe coding features are included in later versions of Dirac. In 2008, BBC transmitted High Definition Television (HDTV) videos of Beijing Olympics using Dirac standard.

CONCLUSION

The recent decades witnessed subsequent and substantial growth in the field of video broadcasting and streaming of real-time video frames over wireless channels. The ease of real-time video streaming also extended its applicability where a user can view online videos on any hand-held devices irrespective of any bandwidth or storage constraints. Video coding determines effective compression on real-time entertainment videos and also collaborate mobile communication scenario to access the modest video frames of specific dimensions. Compression is a reversible conversion (encoding) of data that contains fewer bits. It allows more efficient storage and transmission of the data. Video compression is the process of encoding a video file in such a way that it consumes less space than the original file and is easier to transmit in the network. There are two basic standards of video compression. They are JPEG (Joint Photographic Experts Group) and MPEG (Moving Picture Experts Group). JPEG is single-frame image compression and consist of a minimum implementation (called a baseline system) which all implementation required to support and various extensions for a specific application.

REFERENCES

1. Carrillo, Paula, HariKalva, and Spyros Magliveras. "Compression independent object encryption for ensuring privacy in video surveillance." In Multimedia and Expo, 2008 IEEE International Conference on, pp. 273-276. IEEE, 2008.
2. David R. Bull, Communicating Pictures: A Course in Image and Video Coding, Academic Press, 27-Jun-2014 - Computers - 560 pages
3. G. He, D. Zhou, Y. Li, Z. Chen, "High-Throughput Power-Efficient VLSI Architecture of Fractional Motion Estimation for Ultra-HD HEVC Video Encoding", IEEE Transactions On Very Large Scale Integration (VLSI) Systems, vol. 23, no. 12, December 2015
4. Iain E. Richardson, H.264 and MPEG-4 Video Compression: Video Coding for Nextgeneration Multimedia, John Wiley & Sons, 06-Feb-2004 - Science - 306 pages
5. L. Trzcianowski, "Subjective Assessment for Standard Television Sequences and Videotoms H.264/AVC Video Coding Standard", Journal of Telecommunications and Information technology, 2015
6. S. Dias, M. Siekmann, S. Bosse, H. Schwarz, "Rate-Distortion Optimised Quantisation For HEVC Using Spatial Just Noticeable Distortion", IEEE-European Signal Processing Conference, 2015
7. S. Panayides, M. S. Pattichis, C. P. Loizou, "An Effective Ultrasound Video Communication System Using Despeckle Filtering and HEVC", IEEE Journal of Biomedical and Health Informatics, vol. 19, no. 2, March 2015
8. T. Nguyen and D. Marpe, "Objective Performance Evaluation of the HEVC Main Still Picture Profile", IEEE Transactions On Circuits And Systems For Video Technology, vol. 25, no. 5, May 2015
9. Y. Ye, Y. He, and X. Xiu, "Manipulating Ultra-High Definition Video Traffic", IEEE Computer Society, 2015
10. Y-H Chen, V Sze, "A Deeply Pipelined CABAC Decoder for HEVC Supporting Level 6.2 High-Tier Applications", IEEE Transactions On Circuits And Systems For Video Technology, vol. 25, no. 5, May 2015

Corresponding Author

Maya Chowksey*

Research Scholar, LNCT University, Bhopal